Impact of Small Employers on Income Inequality

John Graves III

April 30, 2021

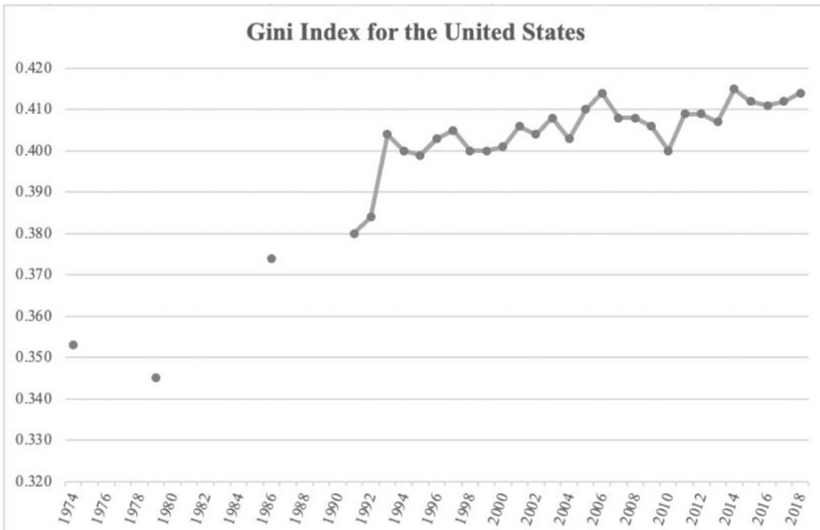
ECON 3161 Dr. Shatakshee Dhongde

Abstract

Income inequality in the United States has increased substantially over the last four decades. This paper attempts to analyze the relationship between the small employers and income inequality (as measured by the Gini index). Other explanatory variables include the natural logarithm of real GDP, percentage of income from retirement income, percentage of income from social security, unemployment rate, percentage of population that identifies as a minority, percentage of population with a high school education, urban population share, and median age. A negative association between income inequality and small employers is hypothesized and supported by the linear regression model in this study.

I. Introduction

It is widely acknowledged by economists, political leaders, and the general populace that income inequality in the United States has increased dramatically over the last few decades. As the graph below represents, the Gini Index, a common measure of inequality, has increased over 20% from since its 1974 level.



World Bank, GINI Index for the United States [SIPOVGINIUSA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/SIPOVGINIUSA, March 22, 2021.

The negative impacts of inequality are not difficult to intuitively understand. Kim, J., & Tebaldi, E. (2013) simplify its importance to two reasons. The first reason is that, according to Berg and Ostry (2011), countries that have more equal distributions of wealth and income tend to also have longer periods of growth. Inequality can contribute to social and political unrest, increasing inefficiency and slowing economic growth. Secondly, income inequality can be viewed negatively as a "social evil" which can have negative effects on the overall happiness of a country.

The purpose of this paper is to assess the size and direction of small employer's impact on income inequality using cross-sectional data and multiple linear regression. The hypothesis is that small employers contribute to a more equal income distribution (negatively impacting the Gini coefficient) and may therefore help combat rising income inequality. The economic and social rationale behind this is not only to contribute to our understanding of income inequality but also to identify a potential solution for local and federal policy to target when attempting to slow the growth of income inequality.

II. Literature Review

In order to study the widely acknowledged growth of income inequality, Hertz and Silva (2020) researched the effect of other income sources on the rising inequality within urban and rural America. The household income data was organized into rural/urban groups using between-group/within-group decomposition to best analyze the impact of differences in average incomes on national income inequality. The research then calculates the impact of each individual income source on inequality. An income source is “regressively distributed” if that income source accounts for a large share of total income, is itself unequally distributed, and it is distributed in a way that reenforces the inequality of the distribution of other income. It will then make a large contribution to overall income inequality and increase the overall Gini coefficient. An income source is “progressively distributed” if lower-income households attribute more of their income to this source and it decreases the overall Gini coefficient. After plotting the contributions of various independent variables to the Gini coefficient, the increasing contributions of retirement and social security income stand out. From 1975-2015 retirement income, social security, and other cash transfers contributed 0.04 to the urban Gini coefficient and 0.05 to the rural coefficient (as large or larger than the contribution of wages and salaries). This identifies urban and rural residential areas, retirement income, and social security income as important contributing factors to wealth inequality.

Hoffmann, Lee, and Lemieux (2020) also examine the contribution of the main explanatory factors behind the growth of income inequality in the last 40 years. Their study finds that a large portion of the growth in income inequality in the 2015-2018 is connected to education. Most of this contribution is derived from between-group and composition effects. The between-group effect is that the widening of the income gap between low-educated and low-educated workers results in a rise in between-group inequality. The composition effect is caused by a group growing in size, which by itself could result in more equality. The study found that increasing returns to education lead to a large increase in the between-group effect and accounted for something near one-third of the variance in income between 2015-2018. A variety of aspects of education appear to be the main force behind the growth of income inequality in the United States.

Wu and Li (2018) completed similar research into additional factors that may contribute to the increasing income inequality in the United States. In their paper, they examine the factors of mobility and volatility that are not traditionally used in existing research. In order to do this analysis in a way that avoided the shortcomings of cross-sectional data, a longitudinal study was carried out using nationally representative data from the Panel Study of Income Dynamics (PSID). The study organized households into three age-controlled cohorts with a family head younger that 45 in the given decades: 1969-1980, 1981-1990, and 1991-2000. The cohorts were then divided into social classes based on income-to-mean ratio. The variation was calculated as the change of income up or down in a given time period. The mobility of the social class was calculated by dividing the average variance by the change in years. The mobility of a social class is the average change-of-status change of status per year for all of the households in that class. Using the Fokker–Planck equation to connect income distributions with the value of variation calculated earlier. The distributions of the mobility of the cohorts in the study show that despite growing income inequality, income mobility has remained constant over the last few decades. However, United States household income volatility has increased significantly since 1990. The study concludes that rising income volatility is a key contributing factor to the rise in income inequality in the United States. In creating a metric opportunity, calculated by the proportion of families in a cohort with increasing income out of all of the families in that cohort, and the decrease in opportunity over time, it is demonstrated that income inequality is positively correlated with volatility.

Each of the referenced papers looks to better understand the causes of income inequality in general and the cause of income inequality in the United States over the last few decades. This paper will look at the impact of small employers, a variable not included in these papers, on income inequality. Calculating the percentage of employees in a county that are employed by a small employer and exploring its relationship with the Gini coefficient through multiple linear regression will provide insight into the ability of small businesses to impact income inequality in the area surrounding them.

III. Data

To analyze and represent the relationship between small employers and income inequality, cross-sectional per county data was gathered from the U.S. Census Bureau, Bureau of Economic Analysis, Bureau of Labor Statistics and the U.S. Labor Department. The dependent variable in this analysis is the Gini coefficient of a given county in the United States, as the Gini coefficient is a summary measure of income inequality. The Gini coefficient equals 0 in the case of perfect equality and 1 in the case of perfect inequality. The 2019 per county Gini coefficient data was sourced from the U.S. Census Bureau. The main explanatory variable used was employees in a county that are employed by an employer that employs less than 20 people (<20 employer), as a percentage of total employees. This variable will sometimes be referred to as the relative size of small business employment. This variable was calculated as the percentage of employees employed by a <20 employer in a county instead of the percentage of <20 employers in a county because this would more accurately represent the size of a county population's income dependence on <20 employers. The 2017 employment data was sourced from the U.S. Department of Labor. It is hypothesized that a larger portion of employees being employed by small businesses will decrease income inequality, giving this variable a negative coefficient. An initial scatterplot of the relationship between the Gini coefficient and relative size of small business employment (Figure 1) shows a mild and negative correlation between the two variables.

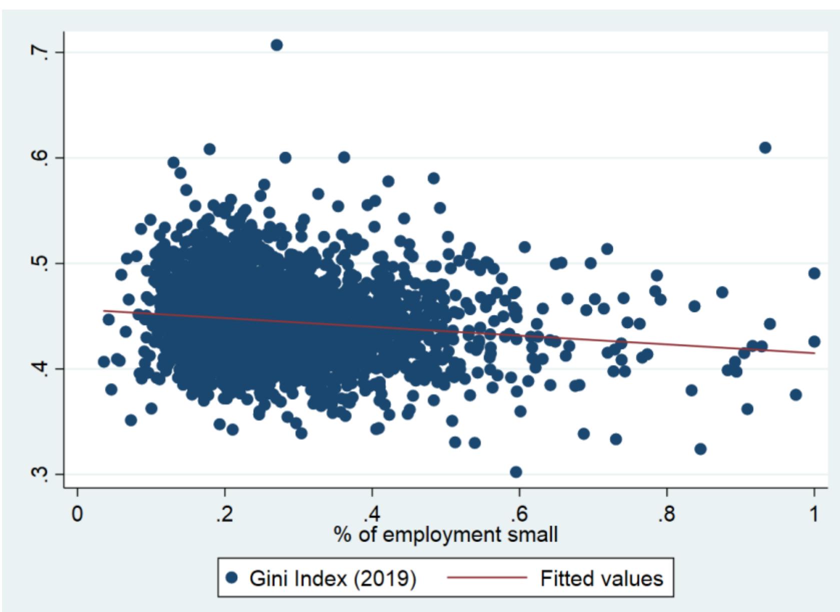


Figure 1 – Scatterplot of Gini coefficient vs. Relative Size of Small Business Employment

In addition to this main variable, the other explanatory variables used were the natural logarithm of real GDP, percentage of income from retirement income, percentage of income from social security, percentage of population that is a minority, percentage of population with a high school education, unemployment, urban population share, and median age. These additional explanatory variables were used to strengthen the multiple linear regression model used to measure the ceteris paribus effect of relative size of small business employment on the Gini coefficient. 2019 data for real GDP was taken from the Bureau of Economic Analysis. 2019 data for unemployment rates was taken from the Bureau of Labor Statistics. 2019 data for percentage of income from retirement income, percentage of income from social security, percentage of population that identifies as a minority, percentage of population with a high school education, urban population share, and median age was taken from the U.S. Census Bureau. A summary of each variable and its variable name is below in Table 1.

Variable Name	Description	Year	Units	Source
gini	Measure of income equality across a population.	2019	Percentage	U.S Census Bureau
smallbiz	Employees in a county that are employed by an employer that employs less than 20 people (<20 employer), as a percentage of total employees	2017	Percentage	Bureau of Labor Statistics
logrealgdp	Natural logirithm of inflation adjusted measure of output	2019	Constant 2019 USD	Bureau of Economic Analysis
retirement	Income that is from retirement income, as a percentage of total income	2019	Percentage	U.S Census Bureau
socialsecur	Income that is from social security, as a percentage of total income	2019	Percentage	U.S Census Bureau
minorities	Population that identifies as a minority group, as a percentage of total population	2019	Percentage	U.S Census Bureau
highschool	Population with the highest academic level as high school, as a percentage of total population	2019	Percentage	U.S Census Bureau
unemployment	Population with highest academic level as bachelor's degree, as a percentage of total population	2019	Percentage	Bureau of Labor Statistics
urban	Population that lives in a desginated urban zone, as a percentage of total population	2019	Percentage	U.S Census Bureau
medage	Median age of total county population	2019	Years	U.S Census Bureau

Table 1 – Variable Descriptions

The natural logarithm of real GDP was used to account for the output of each county adjusted for inflation. A higher *logrealgdp* indicates a county with a larger inflation adjusted economic output. It is hypothesized that a county with higher inflation adjusted economic output will also have higher income inequality given the identified trend of rising inequality, giving this variable a positive coefficient. To account for the portion of a county's income that is retirement income, the variable *retirement* was calculated by dividing the county total retirement income by the county total income. A higher *retirement* explains that a county is more reliant on retirement income. As found by Hertz and Silva (2020), it is hypothesized that a county with higher portion of income being retirement income will experience higher income inequality, giving this variable a positive coefficient. To account for the portion of a county's income that is social security income, the variable *socialsecurity* was calculated by dividing the county total social security income by the county total income. A higher *socialsecurity* explains that a county is more reliant on social security for income. Again, as found by Hertz and Silva (2020), it is hypothesized that a county with higher portion of income being social security will experience higher income inequality, giving this variable a positive coefficient. The percentage of the county population that identifies as a minority is represented by *minorities.* This variable is a measure of the density of minority groups in a county's population. The percentage of the county population with high school as highest academic level of achievement is represented by *highschool*. This variable is used as a measure for the significance of highschool educated workers in the county's workforce. As found by Hoffmann, Lee, and Lemieux (2020), it is hypothesized that a larger level of educated workers will decrease inequality so the coefficient will be negative. The unemployment rate of a county is accounted for by the variable *unemployment.* This variable represents the number of unemployed people in a county as a percentage of the county's labor force. To account for the urban population share of a county, the variable *urban* was created to represent the county population that lives in a designated urban area as a percentage of the total county population. The variable *medage* is used to control for the median age of the population of a county.

The descriptive statistics for each variable can be found below in Table 2.

Variable	Obs	Mean	Std. Dev.	Min	Max
gini	3,075	0.45	0.04	0.30	0.71
smallbiz	3,075	0.27	0.12	0.04	1.00
logrealgdp	3,075	13.97	1.58	10.04	1.00
retirement	3,075	0.11	0.04	0.02	20.40
socialsecur	3,075	0.07	0.03	0.01	0.26
minorities	3,075	0.15	0.16	0.01	0.28
highschool	3,075	0.29	0.06	0.07	0.49
unemployment	3,075	0.04	0.01	0.01	0.21
urban	3,075	0.41	0.31	0.00	1.00
medage	3,075	41.48	5.36	23.50	67.40

Table 2 – Variable Descriptive Statistics

Economic data was gathered for 3,139 counties and county equivalents in the United States, but 64 counties and county equivalents had to be excluded due to insufficient data (list of the excluded is contained in Appendix A).

Before creating each of the regression models in IV. Results, all Classical Linear Model assumptions were checked. Each of the Classical Linear Model assumptions is detailed below:

Assumption 1: The model is linear in parameters so that:

$$y = B_0 + B_1x_1 + B_2x_2 + \ldots + B_kx_k + u$$

All of the models in IV. Results are linear in parameters and satisfy this assumption.

Assumption 2: Random sampling was used in data selection

The data used was sourced from the data sources available and the only excluded data was excluded because of lacking the analyzed variables. There was no method or consideration given to the selection of data, proving that it was random.

Assumption 3: No perfect collinearity in explanatory variables

A table of correlation coefficients between the variables was used to check for perfect collinearity among explanatory variables. Assumption 3 is not violated by any of the explanatory variables because there are no perfectly linear relationships and no constant variables. The table of correlation coefficients can be found in Appendix B.

Assumption 4: Zero conditional mean

The value of the explanatory values in the model must not contain any information about the mean of the unobserved factors. This is difficult to assume so the results should be interpreted with caution.

Assumption 5: Homoskedasticity

The value of the explanatory variables must not contain any information about the variance of the error term. This is also difficult to assume so the results should be interpreted with caution.

Assumption 6: Normality of error terms

The normality of the error term was tested by a plot of the error distribution. The plot shows some indication of non-normality so Assumption 6 may be subject to question. The plot of the error distribution can be found in Appendix C.

IV. Results

Several regression models were created to test the hypothesis. The full STATA regression outputs can be found in Appendix C.

Model 1:

A simple linear regression model to represent the relationship between the Gini coefficient and the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer). The model is written as:

Model 1: $gini = B_0 + B_1(smallbiz) + u$

The number of observations in the model is 3,075 counties and county equivalents. From the STATA regression output, the estimated equation is:

Estimated Equation 1: $gini = 0.4564 - 0.0413(smallbiz)$

Model 1 has an R-squared value of 0.0183, indicating a weak correlation between the Gini coefficient and the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer). The *smallbiz* coefficient has a negative sign, as predicted earlier, and is significant at the 1% level. This means that a 1% increase in the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer) results in a decrease of 0.000413 in the Gini coefficient.

Model 2:

A multiple linear regression that incorporates all variables in order to better explain the variation in the Gini coefficient. The model is written as:

Model 2:

$$gini = B_0 + B_1(smallbiz) + B_2(logrealgdp) + B_3(retirement) + B_4(socialsecur) + B_5(minorities) + B_6(highschool) + B_7(unemployment) + B_8(urban) + B_9(medage)$$

The number of observations in the model is 3,075 counties and county equivalents. From the STATA regression output, the estimated equation is:

Estimated Equation 2:

$$gini = 0.3961 - 0.0165(smallbiz) + 0.0026(logrealgdp) + 0.2174(retirement) - 0.0589(socialsecur) + 0.0709(minorities) – 0.1115(highschool) + 0.3413(unemployment) - 0.00361(urban) + 0.0002(medage)$$

Model 2 has an R-squared value of 0.2114, indicating a moderately weak, but stronger, correlation between Gini coefficient and the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer). The *smallbiz* coefficient -0.0165 is smaller than the coefficient -0.0413 in the simple regression model, indicating that adding explanatory variables reduced the simple regression's overestimation of the impact of the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer) on the Gini coefficient due to omitted variable bias. This means that a 1% increase in the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer) results in a decrease of 0.000165 in the Gini coefficient. *Urban* and *medage* are individually insignificant, *smallbiz* and *socialsecur* are significant at the 5% level, and all of the other variables are significant at the 1% level.

Model 3:

A multiple linear regression that incorporates all variables, excluding *urban* and *medage*, in order to better explain the variation in the Gini coefficient. This model does not incorporate *urban* because *urban* has a p-value of 0.291, indicating that it is not individually significant. This model does not incorporate *medage* because *medage* has a p-value of 0.292, indicating that it is not individually significant. The model is written as:

Model 3:

$$gini = B_0 + B_1(smallbiz) + B_2(logrealgdp) + B_3(retirement) + B_4(socialsecur) + B_5(minorities) + B_6(highschool) + B_7(unemployment)$$

The number of observations in the model is 3,075 counties and county equivalents. From the STATA regression output, the estimated equation is:

Estimated Equation 3:

$$gini = 0.4029 - 0.0122(smallbiz) + 0.0023(logrealgdp) + 0.2318(retirement) - 0.0514(socialsecur) + 0.0699(minorities) - 0.1066(highschool) + 0.3281(unemployment)$$

Model 2 has an R-squared value of 0.2107, indicating a moderately weak correlation between Gini coefficient and the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer) that is not stronger than Estimated Equation 2. This means that a 1% increase in the percentage of employees in a county that are employed by an employer that employs less than 20 people (<20 employer) results in a decrease of 0.000122 in the Gini coefficient. *Smallbiz* and *socialsecur* are significant at the 10% level, and all of the other variables are significant at the 1% level.

Table 3 is a summary of the three linear regression models discussed in the pages above.

Dependent Variable: gini			
Independent Variables	Model 1	Model 2	Model 3
smallbiz	-0.0413*** (0.005)	-0.0165** (0.007)	-0.0122* (0.006)
logrealgdp		0.0026*** (0.001)	0.0023*** (0.001)
retirement		0.2174*** (0.028)	0.2318*** (0.026)
socialsecur		-0.0589** (0.028)	-0.0514* (0.027)
minorities		0.0709*** (0.004)	0.0699*** (0.004)
highschool		-0.1115*** (0.012)	– 0.1066*** (0.011)
unemployment		0.3413*** (0.048)	0.3281*** (0.047)
urban		-0.0036 (0.003)	
medage		0.0002 (0.0002)	
Intercept	0.4564*** (0.002)	0.3961*** (0.012)	0.4029*** (0.011)
Observations	3,075	3,075	3,075
R-squared	0.0183	0.2114	0.2107
Adjusted R-squared	0.018	0.2091	0.2089
			*Significant at 10%, **5%, ***1%

Table 3 - Summary of regressions

V. Extension

<u>F-tests:</u>

In the analysis of Model 2, we found that *urban* and *medage* were individually significant at the 1%, 5%, and 10% levels. To conclude that they are jointly insignificant we must perform a joint F-test on *urban* and *medage.* Our assumptions for this model are:

$$H_0: B_8 = 0, B_9 = 0 \text{ against } H_1: H_0 \text{ is not true}$$

Model 2 will be used as our unrestricted model and Model 3 as our restricted model. The equation that will be used to conduct the F-test is below:

$$\frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k - 1)}$$

The F-value of this test is 1.383. The critical value for the 10% level with 3,075 observations is 2.30. The critical value is larger than the F-value, so we cannot reject H_0 and must conclude that *urban* and *medage* are not jointly significant. This informs us that Model 3 is the most accurate model.

VI. Conclusions

After constructing the three multiple linear regression models above, the original hypothesis is supported. Despite the models not having large R-squared values, the *smallbiz* variable had a negative coefficient and a p-value of <0.1 demonstrating its individual significance. Although the magnitude of the coefficient of *smallbiz* is small, a 1% change in small employer employment decreases the overall Gini coefficient by 0.0122%. This change is even more significant when put into context of the United States Gini coefficient. The overall United States Gini Coefficient is 0.415 and has increased by 0.06 over the last four decades; therefore, a 1% change in small employer employment creating a reduction of 0.000122 is an impact worth noticing. The models all can improve drastically by explaining more of the variance of the Gini coefficient. Every variable used in this research except *urban* and *medage* proved to be individually significant but additional variables like industry concentration could be added to potentially increase the R-squared value. Income inequality is an increasingly pressing issue in the United States and additional research into employers that naturally distribute income more equally could be part of the solution.

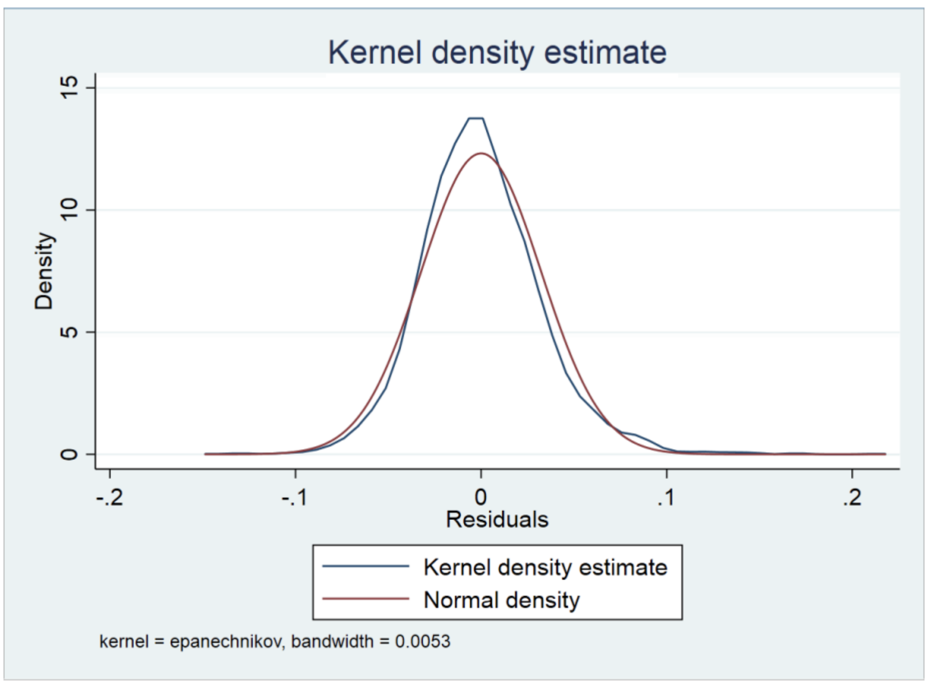
Appendix A. List of U.S. counties and county equivalents excluded from the study due to inadequate data:

Counties excluded due to insufficient data	
Kusilvak Census Area, Alaska	Covington city, Virginia
Oglala Lakota County, South Dakota	Danville city, Virginia
Kenedy County, Texas	Emporia city, Virginia
Prince William County, Virginia	Fairfax city, Virginia
Pulaski County, Virginia	Falls Church city, Virginia
Rappahannock County, Virginia	Franklin city, Virginia
Richmond County, Virginia	Fredericksburg city, Virginia
Roanoke County, Virginia	Galax city, Virginia
Rockbridge County, Virginia	Hampton city, Virginia
Rockingham County, Virginia	Harrisonburg city, Virginia
Russell County, Virginia	Hopewell city, Virginia
Scott County, Virginia	Lexington city, Virginia
Shenandoah County, Virginia	Lynchburg city, Virginia
Smyth County, Virginia	Manassas city, Virginia
Southampton County, Virginia	Manassas Park city, Virginia
Spotsylvania County, Virginia	Martinsville city, Virginia
Stafford County, Virginia	Newport News city, Virginia
Surry County, Virginia	Norfolk city, Virginia
Sussex County, Virginia	Norton city, Virginia
Tazewell County, Virginia	Petersburg city, Virginia
Warren County, Virginia	Poquoson city, Virginia
Washington County, Virginia	Portsmouth city, Virginia
Westmoreland County, Virginia	Radford city, Virginia
Wise County, Virginia	Richmond city, Virginia
Wythe County, Virginia	Roanoke city, Virginia
York County, Virginia	Salem city, Virginia
Alexandria city, Virginia	Staunton city, Virginia
Bristol city, Virginia	Suffolk city, Virginia
Buena Vista city, Virginia	Virginia Beach city, Virginia
Charlottesville city, Virginia	Waynesboro city, Virginia
Chesapeake city, Virginia	Williamsburg city, Virginia
Colonial Heights city, Virginia	Winchester city, Virginia

Appendix B. Correlation coefficients between each variable to satisfy Gauss-Markov Assumption 3

	gini	*smallbiz*	*logrealgdp*	*retirement*	*socialsecur*	*minorities*	*highschool*	*unemployment*	*urban*	*medage*
gini	1.0000									
smallbiz	-0.1354	1.0000								
logrealgdp	0.1167	-0.5858	1.0000							
retirement	0.0964	0.2348	-0.5296	1.0000						
socialsecur	0.1146	0.0767	-0.1395	0.5595	1.0000					
minorities	0.3812	-0.2208	0.1411	-0.0424	0.0422	1.0000				
highschool	-0.1436	0.0857	-0.4375	0.4463	0.0789	-0.1182	1.0000			
unemployment	0.2500	-0.0263	-0.1207	0.3798	0.3492	0.3051	0.1959	1.0000		
urban	0.1069	-0.5464	0.8006	-0.4980	-0.1338	0.1541	-0.4825	-0.0899	1.0000	
medage	-0.0727	0.4517	-0.4053	0.5855	0.4364	-0.3071	0.2934	0.0391	-0.4857	1.0000

Appendix C. Plot of the error distribution to satisfy Classical Linear Model Assumption 6



Appendix D. Regression model outputs

Model 1:

```
. regress gini smallbiz

      Source |       SS           df       MS      Number of obs   =     3,075
-------------+----------------------------------   F(1, 3073)      =     57.40
       Model |  .074949536         1  .074949536   Prob > F        =    0.0000
    Residual |  4.01242493     3,073  .001305703   R-squared       =    0.0183
-------------+----------------------------------   Adj R-squared   =    0.0180
       Total |  4.08737446     3,074   .00132966   Root MSE        =    .03613

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    smallbiz |   -.041317   .0054534    -7.58   0.000    -.0520097   -.0306243
       _cons |   .4564169   .0015874   287.52   0.000     .4533044    .4595294
------------------------------------------------------------------------------
```

Model 2:

```
. regress gini smallbiz logrealgdp retirement socialsecur minorities highschool unemployment urban medage

      Source |       SS           df       MS      Number of obs   =     3,075
-------------+----------------------------------   F(9, 3065)      =     91.31
       Model |  .864173678         9  .096019298   Prob > F        =    0.0000
    Residual |  3.22320078     3,065  .001051615   R-squared       =    0.2114
-------------+----------------------------------   Adj R-squared   =    0.2091
       Total |  4.08737446     3,074   .00132966   Root MSE        =    .03243

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    smallbiz |  -.0165405   .0068991    -2.40   0.017    -.0300678   -.0030132
  logrealgdp |   .0026235   .0006959     3.77   0.000     .0012591     .003988
  retirement |   .2173589   .0279503     7.78   0.000     .1625557    .2721622
 socialsecur |  -.0589496   .0283023    -2.08   0.037     -.114443   -.0034563
  minorities |   .0709094    .004118    17.22   0.000     .0628351    .0789837
  highschool |  -.1115303   .0118051    -9.45   0.000     -.134677   -.0883837
unemployment |   .3413059   .0482406     7.08   0.000     .2467187    .4358931
       urban |  -.0036126   .0034236    -1.06   0.291    -.0103254    .0031001
      medage |   .0001742   .0001653     1.05   0.292      -.00015    .0004984
       _cons |   .3961744   .0118094    33.55   0.000     .3730193    .4193295
------------------------------------------------------------------------------
```

Model 3:

```
. regress gini smallbiz logrealgdp retirement socialsecur minorities highschool unemployment
```

Source	SS	df	MS			
Model	.861265404	7	.123037915	Number of obs	=	3,075
Residual	3.22610906	3,067	.001051878	F(7, 3067)	=	116.97
				Prob > F	=	0.0000
Total	4.08737446	3,074	.00132966	R-squared	=	0.2107
				Adj R-squared	=	0.2089
				Root MSE	=	.03243

gini	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
smallbiz	-.0122284	.0063413	-1.93	0.054	-.0246619	.0002052
logrealgdp	.0022824	.0005651	4.04	0.000	.0011744	.0033903
retirement	.2317878	.0258179	8.98	0.000	.1811657	.2824098
socialsecur	-.0514442	.0272494	-1.89	0.059	-.1048731	.0019847
minorities	.0699017	.0039964	17.49	0.000	.0620658	.0777376
highschool	-.1066113	.0113677	-9.38	0.000	-.1289003	-.0843223
unemployment	.3281402	.0473618	6.93	0.000	.2352761	.4210043
_cons	.402689	.0111149	36.23	0.000	.3808956	.4244824

ACS DEMOGRAPHIC AND HOUSING ESTIMATES. Explore Census Data. (2019). https://data.census.gov/cedsci/table?q=minority+population&tid=ACSDP1Y2019.DP05.

Age and Sex. Explore Census Data. (n.d.). https://data.census.gov/cedsci/table?q=Age+and+Sex&t=Class+of+Worker&g=0100000US.050000_0400000US01.050000&tid=ACSST1Y2019.S0101&hidePreview=true.

ANNUAL ESTIMATES OF THE RESIDENT POPULATION: APRIL 1, 2010 TO JULY 1, 2019. Explore Census Data. (n.d.). https://data.census.gov/cedsci/table?q=urban+population+share&g=0100000US.050000&y=2019&tid=PEPPOP2019.PEPANNRES.

Berg, A. & Ostry, J. (2011). Inequality and unsustainable growth: two sides of the same coin?. IMF Staff Discussion Note, SDN/11/08.

Bureau, U. S. C. (2020, September 15). *Income and Poverty in the United States: 2019*. The United States Census Bureau. https://www.census.gov/data/tables/2020/demo/income-poverty/p60-270.html.

Explore Census Data. (2019). https://data.census.gov/cedsci/table?q=minority+population&tid=ACSDP1Y2019.DP05.

GDP by County, Metro, and Other Areas. U.S. Bureau of Economic Analysis (BEA). (2019). http://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas.

Hertz, T., & Silva, A. (2020). Rurality and Income Inequality in the United States, 1975–2015. Rural Sociology, 85(2), 436–467. https://doi.org/10.1111/ruso.12295

Hoffmann, F., Lee, D. S., & Lemieux, T. (2020). Growing Income Inequality in the United States and Other Advanced Economies. Journal of Economic Perspectives, 34(4), 52–78. https://doi.org/10.1257/jep.34.4.52

Kim, J., & Tebaldi, E. (2013). Trends and sources of income inequality in the united states. *The Journal of Business and Economic Studies*, *19*(2), 1-13,79-80. Retrieved from https://go.openathens.net/redirector/gatech.edu?url=https://search-proquest-com.eu1.proxy.openathens.net/scholarly-journals/trends-sources-income-inequality-united-states/docview/1624963990/se-2?accountid=11107

Pew Hispanic Center. (2008). 2007 Hispanic Healthcare Survey [Data file and code book]. Available from Pew Hispanic Center Web site: http://pewhispanic.org/datasets/

SELECTED ECONOMIC CHARACTERISTICS. Explore Census Data. (2019). https://data.census.gov/cedsci/table?q=retirement+income&t=Class+of+Worker&g=0400000US01.050000&tid=ACSDP1Y2019.DP03&hidePreview=true.

SELECTED POPULATION PROFILE IN THE UNITED STATES. Explore Census Data. (2019). https://data.census.gov/cedsci/table?q=Education&t=Class+of+Worker&g=0100000US.050000_0400000US01.050000&tid=ACSSPP1Y2019.S0201&hidePreview=true.

U.S. Bureau of Labor Statistics. (n.d.). *Local Area Unemployment Statistics Home Page*. U.S. Bureau of Labor Statistics. https://www.bls.gov/lau/.

Wu, Huixuan. (2018). Mobility and volatility: What is behind the rising income inequality in the United States. Physica A, 492.